Selective Coding Strategy for Unicast Composite Networks

Arash Behboodi
Dept. of Telecommunications, SUPELEC
91192 Gif-sur-Yvette, France
Email: {arash.behboodi}@supelec.fr

Pablo Piantanida
Dept. of Telecommunications, SUPELEC
91192 Gif-sur-Yvette, France
Email: {pablo.piantanida}@supelec.fr

Abstract—Consider a composite unicast relay network where the channel statistic is randomly drawn from a set of conditional distributions indexed by $\theta \in \Theta$, which is assumed to be unknown at the source, fully known at the destination and only partly known at the relays. Commonly, the coding strategy at each relay is fixed regardless of its channel measurement. A novel coding for unicast composite networks with multiple relays is introduced. This enables the relays to select dynamically –based on its channel measurement– the best coding scheme between compress-and-forward (CF) and decode-and-forward (DF). As a part of the main result, a generalization of Noisy Network Coding is shown for the case of unicast general networks where the relays are divided between those using DF and CF coding. Furthermore, the relays using DF scheme can exploit the help of those based on CF scheme via offset coding. It is demonstrated via numerical results that this novel coding, referred to as Selective Coding Strategy (SCS), outperforms conventional coding schemes.

I. INTRODUCTION

Multiterminal networks are the essential part of modern telecommunication systems. Recently, these were studied from various aspects. The cutset bound for the general multicast network was established in [1]. Network coding theorem for the graphical multicast network was investigated in [2] where the max-flow min-cut theorem for network information flow was presented for the point-to-point communication network. Whereas Lim *et al.* proposed the Noisy Network Coding (NNC) scheme for the general multicast network, which includes most of the existing bounds on multiterminal networks [3]. Kramer *et al.* developed an inner bound for a point-to-point general network using decode-and-forward (DF) which achieves the capacity of the degraded multicast network [4].

For the above mentioned scenarios, the probability distribution (PD) of the network is supposed to be fixed during the communication and hence available to all nodes beforehand. However, wireless channels are essentially time-varying due to fading and user mobility, and hence the terminals do not have full knowledge of all channel parameters involved in the communication. In particular, without feedback channel state information (CSI) cannot be available to the encoders end. During years, an ensemble of research activities has been dedicated to both theoretical and practical aspects of communication in presence of channel uncertainty. From an information-theoretic viewpoint, the compound channel, first

introduced by Wolfowitz [5] is one of the most important models that deals with channel uncertainty, and continues to attract much attention from researchers (see [6] and references therein). Composite models are more appropriate to deal with wireless scenarios since unlike compound models they deal with channel uncertainty by introducing a PD $\mathbb{P}_\theta$ on the channel selection. These models consist of a set of conditional PDs from which the current channel index θ, which can be a vector of parameters, is drawn according to $\mathbb{P}_\theta$ and remains fixed during the communication. An example of this model can be slowly fading channels. Capacity for this class of channels has been widely studied beforehand (see [7] and references therein), for wireless scenarios via the well-known notion of outage capacity (see [8] and references therein) and oblivious cooperation over fading Gaussian channels in [9]–[11].

In this paper, we study the composite multiple relay network where the channel index $\theta \in \Theta$ is randomly drawn according to $\mathbb{P}_\theta$. The index $\theta = (\theta_r, \theta_d)$ remains fixed during the communication but is unknown at the source, fully known at the destination and partly known θ_r at the relays end. Although a compound approach can guarantee asymptotically zero-error probability regardless of θ, it would be not an adequate choice for most of wireless models. As a different approach, the coding rate r is selected regardless of the current index. Hence the encoder cannot necessarily guarantee arbitrary small error probability. In this case the asymptotic error probability becomes the measure of interest, characterizing the reliability function. Moreover, it turns out that depending on the channel draw, there may not be a unique set of relay functions that minimizes the error probability. In other words, the relay function should be chosen based on the channel parameters. However, since full CSI is not available to all nodes, the relay functions are usually chosen regardless of their channel measurements which becomes the bottleneck of the coding. We present a novel coding strategy from which the relays can select, based on their measurements, an adequate coding strategy. To this purpose an achievable region that generalizes NNC to the case of mixed coding strategy, where DF relays exploit the help of CF relays, is derived.

Section II presents definitions and Section III introduces main results, and the sketch of proofs is relegated to Section IV. Finally, numerical evaluation over the slow-fading Gaussian two relay channel is presented in Section V.
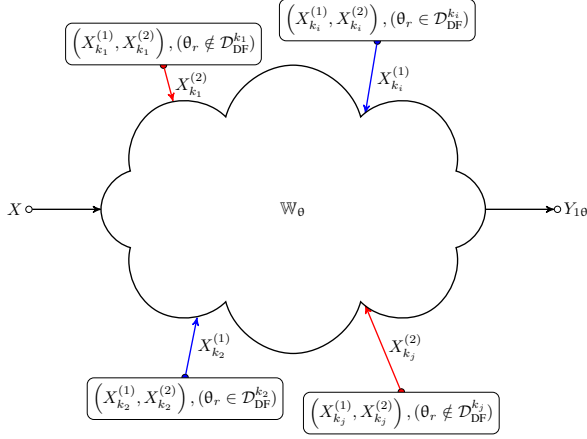
Fig. 1. Composite Multiple Relay Network.

II. PROBLEM DEFINITION

The composite multiple relay channel consists of a set of multiple relay channels, as depicted in Fig. 1 and denoted by

$$\left\{ \mathbb{W}_\theta^n = P_{Y_{1,\theta}^n Z_{1\theta_r}^n Z_{2\theta_r}^n \cdots Z_{N\theta_r}^n | X^n X_{1\theta_r}^n X_{2\theta_r}^n \cdots X_{N\theta_r}^n} \right\}_{n=1}^\infty$$

where X denotes the channel input, $X_{k\theta_r}$ and $Z_{k\theta_r}$ the relay inputs and outputs and $Y_{1\theta}$ the channel output. We assume a memoryless multiple relay channel with N relays but single source and destination. The channels are indexed by vector of parameters $\theta = (\theta_d, \theta_r)$ with $\theta_d \in \Theta_d, \theta_r \in \Theta_r$, where θ_r denotes all parameters affecting the relays' output and θ_d are the remaining parameters involved in the communication. Let $\mathbb{P}_\theta$ be a joint probability measure on Θ. The channel parameters affecting relay and destination outputs $\theta = (\theta_r, \theta_d)$ are drawn according to the joint PD $\mathbb{P}_\theta$ and remain fixed during the communication. However, the specific draw of θ is assumed to be unknown at the source, fully known at the destination and partly known θ_r at the relays end. Assume that $\mathcal{N} = \{1, \ldots, N\}$ and for any $\mathcal{S} \subseteq \mathcal{N}$, $X_{\mathcal{S}} = \{X_i : i \in \mathcal{S}\}$.

Definition 1 (code and achievability): A code-$\mathcal{C}(n, M_n, r)$ for the composite multiple relay channel consists of:

- An encoder mapping $\{\varphi : \mathcal{M}_n \longmapsto \mathcal{X}^n\}$,
- A decoder mapping $\{\phi_\theta : \mathcal{Y}_1^n \longmapsto \mathcal{M}_n\}$,
- A set of relay functions $\left\{ f_{i\theta_r}^{(k)} : \mathcal{Z}_k^{i-1} \longmapsto \mathcal{X}_k \right\}_{i=1}^n$ for $k \in \mathcal{N}$. Only partial CSI at the relay is assumed which is mainly related to the k-th source-to-relay channel.

An error probability $0 \leq \epsilon < 1$ is said to be r-achievable, if there exists a code-$\mathcal{C}(n, M_n, r)$ with rate satisfying

$$\liminf_{n \to \infty} \frac{1}{n} \log M_n \geq r \qquad (1)$$

and average error probability

$$\limsup_{n \to \infty} \mathbb{E}_\theta \left[\Pr \left\{ \phi_\theta(Y_{1\theta}^n) \neq W \middle| \theta \right\} \right] \leq \epsilon. \qquad (2)$$

The infimum of all r-achievable EPs $\bar{\epsilon}(r)$ is defined as

$$\bar{\epsilon}(r) = \inf \{0 \leq \epsilon < 1 : \epsilon \text{ is } r\text{-achievable}\}. \qquad (3)$$

We emphasize that for channels satisfying the strong converse property and with unique best code word, e.g., Gaussian slow-fading single user channel, (3) coincides with the definition of the outage probability for the unique best codeword.

In the present setting, we assume that the source is not aware of the specific draw $\theta \sim \mathbb{P}_\theta$ and hence, the coding rate r and the coding strategy –DF or CF scheme– must be chosen independent of the channel draw. Furthermore, both remain fixed during the communication regardless of the channel measurement at the relays end. We aim to characterize the smallest possible average error probability as defined by (2), as a function of the coding rate r.

It can be shown that $\bar{\epsilon}(r)$ can be bounded as follows

$$\mathbb{P}_\theta(r \in \mathcal{S}_\theta) \leq \bar{\epsilon}(r) \leq \inf_{\mathcal{C}} \mathbb{P}_\theta(r \notin \mathcal{R}_\theta(\mathcal{C})), \qquad (4)$$

where $\mathcal{R}_\theta$ is any achievable rate for the unicast network with a given θ, and $\mathcal{S}_\theta$ is the infimum of all rates such that every code with such rate yields error probability tending to one, and $\mathcal{C}$ as all codes. It can be shown that S_θ can be replaced with max-flow min-cut bound. A special case of this result has been proved in recent work [11] for the relay channel.

III. COMPOSITE MULTIPLE RELAY NETWORKS

Consider the composite unicast network with multiple relays and parameters θ. The rate is fixed to r and so is the source code. The goal is to minimize the expected error probability. The common option is that each relay fixes its coding strategy, namely DF or Compress-and-Forward (CF), regardless of θ. In other words, the relays with index in $\mathcal{V} \subseteq \mathcal{N}$ will use CF scheme. For instance, to evaluate the expected error probability we first present an achievable rate for the multiple relay network where NNC is generalized to networks with mixed cooperative strategy. Part of relays are using DF coding while the reminding relays use CF scheme. Moreover, DF relays exploit the help of CF relays to decode the source message. Using this theorem, an achievable rate can be obtained for every set $\mathcal{V}$ of relay nodes.

Theorem 1 (Cooperative Mixed NNC): For the multiple relay channel, the following rate is achievable

$$R \leq \max_{P \in \mathcal{P}} \max_{\mathcal{V} \subseteq \mathcal{N}} \min \left\{ \max_{\mathcal{T} \in \Upsilon(\mathcal{V})} \min_{\mathcal{S} \subseteq \mathcal{T}} R_\mathcal{T}(\mathcal{S}), \right.$$
$$\left. \min_{k \in \mathcal{V}^c} \max_{\mathcal{T}_k \in \Upsilon_k(\mathcal{V})} \min_{\mathcal{S} \subseteq \mathcal{T}_k} R_{\mathcal{T}_k}^{(k)}(\mathcal{S}) \right\} \qquad (5)$$

with

$$R_\mathcal{T}(\mathcal{S}) = I(XX_{\mathcal{V}^c}X_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c}Y_1 | X_{\mathcal{S}^c}Q)$$
$$\qquad - I(Z_{\mathcal{S}}; \hat{Z}_{\mathcal{S}} | XX_{\mathcal{T} \cup \mathcal{V}^c}\hat{Z}_{\mathcal{S}^c}Y_1Q) \qquad (\mathcal{S}^c = \mathcal{T} - \mathcal{S}),$$

$$R_{\mathcal{T}_k}^{(k)}(\mathcal{S}) = I(X; \hat{Z}_{\mathcal{T}_k}Z_k | X_{\mathcal{V}^c}X_{\mathcal{T}_k}Q) + I(X_{\mathcal{S}}; Z_k | X_{\mathcal{V}^c \cup \mathcal{S}^c}Q)$$
$$\qquad - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{V}^c \cup \mathcal{T}_k}\hat{Z}_{\mathcal{S}^c}Z_kQ) \qquad (\mathcal{S}^c = \mathcal{T}_k - \mathcal{S}),$$

for $\mathcal{T}, \mathcal{T}_k \subseteq \mathcal{V} \subseteq \mathcal{N}$ and $\mathcal{V}^c = \mathcal{N} - \mathcal{V}$. Moreover $\Upsilon(\mathcal{V})$ and $\Upsilon_k(\mathcal{V})$ are defined by

$$\Upsilon(\mathcal{V}) = \{\mathcal{T} \subseteq \mathcal{V} : \text{ for all } \mathcal{S} \subseteq \mathcal{T}, Q_\mathcal{T}(\mathcal{S}) \geq 0\}, \qquad (6)$$
$$\Upsilon_k(\mathcal{V}) = \{\mathcal{T} \subseteq \mathcal{V} : \text{ for all } \mathcal{S} \subseteq \mathcal{T}, Q_\mathcal{T}^{(k)}(\mathcal{S}) \geq 0\}, \qquad (7)$$

while $Q_{\mathcal{T}}(\mathcal{S})$ and $Q_{\mathcal{T}}^{(k)}(\mathcal{S})$ are given by

$$Q_{\mathcal{T}}(\mathcal{S}) = I(X_{\mathcal{S}} \; ; \hat{Z}_{\mathcal{S}^c} Y_1 | XX_{\mathcal{S}^c \cup \mathcal{V}^c} Q)$$
$$-I(Z_{\mathcal{S}}; \hat{Z}_{\mathcal{S}} | XX_{\mathcal{T} \cup \mathcal{V}^c} \hat{Z}_{\mathcal{S}^c} Y_1 Q), \quad (8)$$

$$Q_{\mathcal{T}}^{(k)}(\mathcal{S}) = I(X_{\mathcal{S}}; Z_k | X_{\mathcal{V}^c \cup \mathcal{S}^c} Q)$$
$$-I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | XX_{\mathcal{V}^c \cup \mathcal{T}} \hat{Z}_{\mathcal{S}^c} Z_k Q). \quad (9)$$

And $\mathcal{P}$ is the set of all admissible distributions:
$$\mathcal{P} = \Big\{ P_{QXX_{\mathcal{N}} Z_{\mathcal{N}} \hat{Z}_{\mathcal{N}} Y_1} = P_Q P_{XX_{\mathcal{V}^c} | Q} P_{Y_1 Z_{\mathcal{N}} | XX_{\mathcal{N}}} \times$$
$$\prod_{j \in \mathcal{V}} P_{X_j | Q} P_{\hat{Z}_j | X_j Z_j Q} \Big\}.$$

Notice that $R_{\mathcal{T}}(\mathcal{S})$ (condition of correct decoding at the destination) is in general better than $R_{\mathcal{T}_k}^{(k)}(\mathcal{S})$ (condition of correct decoding at the relay k). This is because the destination uses backward decoding. Moreover using the same technique as [12], it can be shown that the optimization in (5) can be done over $\mathcal{T} \subseteq \mathcal{V}$ instead of $\mathcal{T} \in \Upsilon(\mathcal{V})$. In other words the relays in $\mathcal{T} \subseteq \mathcal{V}^c$ can increase the rate only if they satisfy (7). It can be seen that for each $\mathcal{T}$, if there is $\mathcal{A} \subseteq \mathcal{T}$ such that $Q_{\mathcal{T}}(\mathcal{A}) < 0$, then we can remove the relays in $\mathcal{A}$ from $\mathcal{T}$ and the rate is improved by not using their compression, which is easier to manipulate in composite setting. By choosing $\mathcal{V} = \mathcal{N}$, the region of Theorem 1 is reduced to the same region, as in [12], [13], which is equivalent to NNC region [3]. So Theorem 1 generalizes and includes the previous NNC scheme and it provides a potentially larger region. Indeed, for the degraded single relay channel, it is capacity achieving while NNC is strictly suboptimal. In fact, relay nodes are divided into two groups. The first group is $\mathcal{V}^c$ which are using DF coding and the second group $\mathcal{V}$ which are using CF scheme.

Now consider the composite setting again. Before starting the communication, the source knows that the relays with index in $\mathcal{V}$ use CF while the others use DF scheme. For each θ and $\mathcal{V}$, the rate r is achievable if it belongs to the region in the previous theorem and otherwise an error is declared. All that CF relays can do in this case is to choose their distribution based on θ_r such that it minimizes the error probability. Thus the expected error probability for the composite multiple relay channel with partial CSI θ_r at the relays is bounded by

$$\bar{\epsilon}(r) \leq \inf_{\mathcal{V} \subseteq \mathcal{N}} \min_{p(x,x_{\mathcal{V}^c},q)}$$

$$\mathbb{E}_{\theta_r} \Big\{ \min_{\prod_{j \in \mathcal{V}} p(x_j | q) p(\hat{z}_j | x_j z_j q)} \mathbb{P}_{\theta | \theta_r} \big[r > I_{\text{CMNNC}}(\mathcal{V}) | \theta_r \big] \Big\}$$

where for all $\theta = (\theta_r, \theta_d)$, $I_{\text{MNNC}}(\mathcal{V}) =$

$$\min \Big\{ \max_{\mathcal{T} \subseteq \mathcal{V}} \min_{\mathcal{S} \subseteq \mathcal{T}} R_{\mathcal{T}}(\mathcal{S}, \theta), \min_{k \in \mathcal{V}^c} \max_{\mathcal{T}_k \in \Upsilon_k(\mathcal{V})} \min_{\mathcal{S} \subseteq \mathcal{T}_k} R_{\mathcal{T}_k}^{(k)}, (\mathcal{S}, \theta) \Big\},$$
$$(10)$$

$$R_{\mathcal{T}}(\mathcal{S}, \theta) = I(XX_{\mathcal{V}^c} X_{\mathcal{S}}; \hat{Z}_{\mathcal{S}^c} Y_{1\theta} | X_{\mathcal{S}^c} Q)$$
$$- I(Z_{\mathcal{S}\theta_r}; \hat{Z}_{\mathcal{S}} | XX_{\mathcal{T} \cup \mathcal{V}^c} \hat{Z}_{\mathcal{S}^c} Y_{1\theta} Q)(\mathcal{S}^c = \mathcal{T} - \mathcal{S}),$$

$$R_{\mathcal{T}_k}^{(k)}(\mathcal{S}, \theta) = I(X; \hat{Z}_{\mathcal{T}_k} Z_{k\theta_r} | X_{\mathcal{V}^c} X_{\mathcal{T}_k} Q)$$
$$+ I(X_{\mathcal{S}}; Z_{k\theta_r} | X_{\mathcal{V}^c \cup \mathcal{S}^c} Q) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}\theta_r} | X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_{k\theta_r} Q)$$

for $\mathcal{T}, \mathcal{T}_k \subseteq \mathcal{V} \subseteq \mathcal{N}$ and $\mathcal{V}^c = \mathcal{N} - \mathcal{V}$. Similarly $\Upsilon_k(\mathcal{V})$ is

$$\Upsilon_k(\mathcal{V}) = \{ \mathcal{T} \subseteq \mathcal{V} : \text{ for all } \mathcal{S} \subseteq \mathcal{T}, Q_{\mathcal{T}}^{(k)}(\mathcal{S}, \theta_r) \geq 0 \} \quad (11)$$

where $Q_{\mathcal{T}}^{(k)}(\mathcal{S}, \theta_r)$ is defined as follows:

$$Q_{\mathcal{T}}^{(k)}(\mathcal{S}, \theta_r) = I(X_{\mathcal{S}}; Z_{k\theta_r} | X_{\mathcal{V}^c \cup \mathcal{S}^c} Q)$$
$$- I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}\theta_r} | XX_{\mathcal{V}^c \cup \mathcal{T}} \hat{Z}_{\mathcal{S}^c} Z_{k\theta_r} Q).$$

In the preceding scheme all relays fix coding regardless of the available CSI. However, it is possible that the relays select and change their coding based on its CSI. To this purpose each relay generates many codebooks and sends one of them which fits the best to the channel with the parameter θ_r. More precisely, each relay k has a decision region $\mathcal{D}_{\text{DF}}^{(k)}$ such that for all $\theta_r \in \mathcal{D}_{\text{DF}}^{(k)}$, the relay k uses DF scheme and otherwise it uses CF scheme. For each $\mathcal{V} \subseteq \mathcal{N}$, define $\mathcal{D}_{\mathcal{V}}$ as follows:

$$\mathcal{D}_{\mathcal{V}} = \left(\bigcap_{k \in \mathcal{V}^c} \mathcal{D}_{\text{DF}}^{(k)} \right) \bigcap \left(\bigcap_{k \in \mathcal{V}} \mathcal{D}_{\text{DF}}^{(k)^c} \right).$$

If $\theta_r \in \mathcal{D}_{\mathcal{V}}$, then $\theta_r \notin \mathcal{D}_{\text{DF}}^{(k)}$ for all $k \in \mathcal{V}$, and $\theta_r \in \mathcal{D}_{\text{DF}}^{(k)}$ for all $k \notin \mathcal{V}$. So the k-th relay, for each $k \in \mathcal{V}$ uses CF and the relay k' for $k' \in \mathcal{V}^c$ uses DF. The ensemble of decision regions of relays will thus provide the regions $\mathcal{D}_{\mathcal{V}}$ which are mutually disjoint and all together form a partitioning over the set Θ_r. Now if $\theta_r \in \mathcal{D}_{\mathcal{V}}$, we have a multiple relay network where the relays in $\mathcal{V}$ are using CF. The achievable rate corresponding to this case is known from Theorem 1.

As shown in Fig. 1, each relay has two set of codewords: $X_{(k)}^{(1)}$ and $X_{(k)}^{(2)}$. The first code $X_{(k)}^{(1)}$ is transmitted when $\theta_r \in \mathcal{D}_{\text{DF}}^{(k)}$. This code is based on DF strategy so the k-th relay decodes the source message and transmits it to the destination. However the source, not knowing whether the k-th relay is sending $X_{(k)}^{(1)}$ or not, uses superposition coding and superimpose its code over $X_{(k)}^{(1)}$. If the k-th relay sends $X_{(k)}^{(1)}$ then this will become DF relaying. The source, oblivious to the relays decision, generates its own code by using superposition coding and X is superimposed over $X_{\mathcal{N}}^{(1)}$, i.e., all possible DF relay inputs. This does not affect $R_{\mathcal{T}}(\mathcal{S}, \theta)$ by applying the proper Markov chain, but changes $R_{\mathcal{T}_k}^{(k)}(\mathcal{S}, \theta)$, as we will see.

On the other hand, if $\theta_r \notin \mathcal{D}_{\text{DF}}^{(k)}$ then CF scheme is used. Note that unlike DF, the code which is used for CF $X_{(k)}^{(2)}$, is independent of the source code and so its PD can be chosen adaptively based on θ_r. The optimum choice for $\mathcal{D}_{\mathcal{V}}$ will potentially give a better outage probability than the case that each relay is using a fixed coding for all θ_r. This provides a non-formal proof of the next theorem.

Theorem 2 (SCS with partial CSI-cooperative relays): The average error probability of the composite multiple relay channel with partial CSI θ_r at the relays can be upper bounded by $\bar{\epsilon}(r) \leq \min_{p(x,x_{\mathcal{N}}^{(1)},q)} \inf_{\{\mathcal{D}_{\mathcal{V}}, \mathcal{V} \subseteq \mathcal{N}\} \in \Pi(\Theta_r, N)}$

$$\sum_{\mathcal{V} \subseteq \mathcal{N}} \mathbb{E}_{\theta_r} \Big\{ \min_{\prod_{j \in \mathcal{V}} p(x_j^{(2)} | q) p(\hat{z}_j | x_j^{(2)} z_j q)}$$
$$\mathbb{P}_{\theta | \theta_r} \big[r > I_{\text{CMNNC}}(\mathcal{V}), \theta_r \in \mathcal{D}_{\mathcal{V}} | \theta_r \big] \Big\}, \quad (12)$$

$\Pi\left(\Theta_r, N\right)$ is the set of all partitioning over Θ_r into at most 2^N disjoint sets. The relay inputs X_k is chosen from $\left(X_k^{(1)}, X_k^{(2)}\right)$ such that X_k is equal to $X_k^{(1)}$ if $\theta_r \in \mathcal{D}_{\mathrm{DF}}^k$ and equal to $X_k^{(2)}$ if $\theta_r \notin \mathcal{D}_{\mathrm{DF}}^k$. Indeed, for $\theta_r \in \mathcal{D}_{\mathcal{V}}$ the next Markov chain holds:

$$\left(X_{\mathcal{V}}^{(1)}, X_{\mathcal{V}^c}^{(2)}\right) \ominus \left(X, X_{\mathcal{V}^c}^{(1)}, X_{\mathcal{V}}^{(2)}\right) \ominus \left(Y_{1\theta}, Z_{\mathcal{N}\theta_r}\right),$$

where $I_{\mathrm{MNNC}}(\mathcal{V})$ and $\Upsilon_k(\mathcal{V})$ are defined by expressions (10) and (11) with the difference that in $R_{\mathcal{T}_k}^{(k)}(\mathcal{S}, \theta)$ and $Q_{\mathcal{T}}^{(k)}(\mathcal{S}, \theta_r)$, $X_{\mathcal{V}^c}$ is replaced with $X_{\mathcal{N}}^{(1)}$.

IV. SKETCH OF THE PROOF OF THEOREM 1

Consider first the two relay network. Relay 1 uses DF scheme to help the source so it has to decode the source messages successively and not backwardly, and Relay 2 uses CF scheme. However, relay 1 wants to exploit the help of relay 2 to decode the source message. So it does not start decoding until it retrieves the compression index. To this end, relay 1 uses offset decoding which means that it waits two blocks instead of one to decode the source message and the compression index. In block $b = 2$, the relay 1 decodes the compression index l_1 and the message w_1. Equally, the source code at block $b + 2$ is correlated with relay 1 code from the block b and not block $b+1$. This comes at the expense of one block of delay. The source has to wait until $b = B + L$ to start backward-decoding. The compression index l_{B+2} is repeated until the block $B + L$. Fix $P, \mathcal{V}, \mathcal{T}$ and $\mathcal{T}_k$'s such that they maximize the right hand side of (5). Assume a set $\mathcal{M}_n$ of size 2^{nR} of message indices W to be transmitted, again in $B + L$ blocks, each of them of length n. At the last $L - 2$ blocks, the last compression index is first decoded and then all compression indices and transmitted messages are jointly decoded. Relays in $\mathcal{V}^c$ start to decode after block 2.

Code generation:

(i) Randomly and independently generate 2^{nR} sequences $\underline{x}_{\mathcal{V}^c}$ drawn i.i.d. from $P_{X_{\mathcal{V}^c}}^n(\underline{x}_{\mathcal{V}^c}) = \prod\limits_{j=1}^{n} P_{X_{\mathcal{V}^c}}(x_{\mathcal{V}^c j})$. Index them as $\underline{x}_{\mathcal{V}^c}(r)$ with index $r \in \left[1, 2^{nR}\right]$.

(ii) For each $\underline{x}_{\mathcal{V}^c}(r)$, randomly and conditionally independently generate 2^{nR} sequences $\underline{x}$ drawn i.i.d. from $P_{X|X_{\mathcal{V}^c}}^n(\underline{x}|\underline{x}_{\mathcal{V}^c}(r)) = \prod\limits_{j=1}^{n} P_{X|X_{\mathcal{V}^c}}(x_j|x_{\mathcal{V}^c j})$. Index them as $\underline{x}(r, w)$, where $w \in \left[1, 2^{nR}\right]$.

(iii) For each $k \in \mathcal{V}$, randomly and independently generate $2^{n\hat{R}_k}$ sequences $\underline{x}_k$ drawn i.i.d. from $P_{X_k}^n(\underline{x}_k) = \prod\limits_{j=1}^{n} P_{X_k}(x_{kj})$. Index them as $\underline{x}_k(r_k)$, where $r_k \in \left[1, 2^{n\hat{R}_k}\right]$ for $\hat{R}_k = I(Z_k; \hat{Z}_k | X_k) + \epsilon$.

(iv) For each $k \in \mathcal{V}$ and each $\underline{x}_k(r_k)$, randomly and conditionally independently generate $2^{n\hat{R}_k}$ sequences $\underline{\hat{z}}_k$ each with probability $P_{\hat{Z}_k|X_k}^n(\underline{\hat{z}}_k|\underline{x}_k(r_k)) = \prod\limits_{j=1}^{n} P_{\hat{Z}_k|X_k}(\hat{z}_{kj}|x_{kj}(r_k))$. Index them as $\underline{\hat{z}}_k(r_k, \hat{s}_k)$, where $\hat{s}_k \in [1, 2^{n\hat{R}_k}]$.

Encoding part:

(i) In every block $i = [1 : B]$, the source sends w_i using $\underline{x}\left(w_{(i-2)}, w_i\right)$ $(w_0 = w_{-1} = 1)$. Moreover, for blocks $i = [B+1 : B+L]$, the source sends the dummy message $w_i = 1$ known to all users.

(ii) For every block $i = [1 : B + L]$, and each $k \in \mathcal{V}^c$, the relay k knows $w_{(i-2)}$ by assumption and $w_0 = w_{-1} = 1$, so it sends $\underline{x}_k\left(w_{(i-2)}\right)$.

(iii) For each $i = [1 : B + 2]$, each $k \in \mathcal{V}$, the relay k after receiving $\underline{z}_k(i)$, searches for at least one index l_{ki} with $l_{k0} = 1$ such that

$$\left(\underline{x}_k(l_{k(i-1)}), \underline{z}_k(i), \underline{\hat{z}}_k(l_{k(i-1)}, l_{ki})\right) \in \mathcal{A}_\epsilon^n[X_k Z_k \hat{Z}_k].$$

The probability of finding such l_{ki} goes to one as n goes to infinity due to the choice of $\hat{R}_k$.

(iv) For $i = [1 : B + 2]$ and $k \in \mathcal{V}$, relay k knows from the previous block $l_{k(i-1)}$ and it sends $\underline{x}_k(l_{k(i-1)})$. Moreover, relay k repeats $l_{k(B+2)}$ for $i = [B+3 : B+L]$, which means for $L - 2$ blocks.

Decoding part:

(i) After the transmission of the block $i = [1 : B + 1]$ and for each $k \in \mathcal{V}^c$, with the assumption that all messages and compression indices up to block $i - 1$ have been correctly decoded, the k-th relay searches for the unique index $(\hat{w}_b, \hat{l}_{\mathcal{T}_k b})$ by looking at two consecutive blocks b and $b + 1$ such that:

$$\Big(\underline{x}(w_{(b-2)}, \hat{w}_b), \underline{x}_{\mathcal{V}^c}(w_{(b-2)}), \underline{z}_k(b), \Big(\underline{x}_k(l_{k(b-1)}),$$
$$\underline{\hat{z}}_k(l_{k(b-1)}, \hat{l}_{kb})\Big)_{k \in \mathcal{T}_k}\Big) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T}_k \cup \mathcal{V}^c} \hat{Z}_{\mathcal{T}_k} Z_k] \text{ and}$$
$$\Big(\underline{x}_{\mathcal{V}^c}(w_{(b-1)}), \underline{z}_k(b+1), (\underline{x}_k(\hat{l}_{kb}))_{k \in \mathcal{T}_k}\Big) \in \mathcal{A}_\epsilon^n[X_{\mathcal{T}_k \cup \mathcal{V}^c} Z_k].$$

Probability of error goes to zero as $n \to \infty$ if

$$R \leq I(X; \hat{Z}_{\mathcal{T}} Z_k | X_{\mathcal{V}^c} X_{\mathcal{T}_k}) + I(X_{\mathcal{S}}; Z_k | X_{\mathcal{V}^c} X_{\mathcal{S}^c})$$
$$- I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k). \tag{13}$$
$$0 \leq I(Z_k; X_{\mathcal{S}} | X_{\mathcal{V}^c \cup \mathcal{S}^c}) - I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | XX_{\mathcal{V}^c \cup \mathcal{T}_k} \hat{Z}_{\mathcal{S}^c} Z_k). \tag{14}$$

Given the fact that $\mathcal{T}_k \in \Upsilon_k(\mathcal{V})$, the last inequality holds for each $\mathcal{S} \subseteq \mathcal{T}_k$.

(ii) The destination jointly searches for the unique indices $(\hat{l}_{k(B+2)})_{k \in \mathcal{T}}$ such that for all $b \in [B + 3 : B + L]$, $\Big((\underline{x}_k(\hat{l}_{k(B+1)}))_{k \in \mathcal{T}}, \underline{x}(1, 1), \underline{x}_{\mathcal{V}^c}(1), \underline{y}_1(b)\Big)$ belongs to $\mathcal{A}_\epsilon^n[XX_{\mathcal{T}} X_{\mathcal{V}^c} Y_1]$. The probability of error goes to zero as $n \to \infty$ provided for all subsets $\mathcal{S} \subseteq \mathcal{T}$:

$$\sum_{k \in \mathcal{S}} I(\hat{Z}_k; Z_k | X_k) \leq (L - 2) I(X_{\mathcal{S}}; XX_{\mathcal{S}^c \cup \mathcal{V}^c} Y_1).$$

(iii) With the assumption that $(w_{b+2}, l_{\mathcal{T}(b+2)})$ have been correctly decoded, the destination finds backwardly the unique pair of indices $(\hat{w}_b, \hat{l}_{\mathcal{T}(b+1)})$ such that:

$$\Big(\underline{x}(\hat{w}_b, w_{(b+2)}), \underline{x}_{\mathcal{V}^c}(\hat{w}_b), \underline{y}_1(b + 2), \big(\underline{x}_k(\hat{l}_{k(b+1)}),$$
$$\underline{\hat{z}}_k(\hat{l}_{k(b+1)}, l_{k(b+2)})\big)_{k \in \mathcal{T}}\Big) \in \mathcal{A}_\epsilon^n[XX_{\mathcal{T} \cup \mathcal{V}^c} \hat{Z}_{\mathcal{T}} Y_1].$$
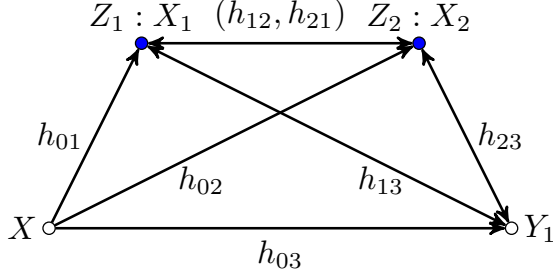
Fig. 2. Fading Gaussian two-relay channel.

The probability of error goes to zero as $n \to \infty$ if:

$$0 < I(\hat{Z}_{\mathcal{S}^c} Y_1; X_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{S}^c})$$
$$- I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1),$$
$$R \leq I(X X_{\mathcal{V}^c} X_{\mathcal{S}}; Y_1 \hat{Z}_{\mathcal{S}^c} | X_{\mathcal{S}^c})$$
$$- I(\hat{Z}_{\mathcal{S}}; Z_{\mathcal{S}} | X X_{\mathcal{V}^c \cup \mathcal{T}} \hat{Z}_{\mathcal{S}^c} Y_1).$$

Using the previous inequalities, and by choosing finite L but large enough, by letting $(B, n) \to \infty$ and adding the time-sharing random variable Q the proof is finished.

V. GAUSSIAN SLOW-FADING NETWORKS

Consider the Gaussian fading two-relay network, depicted in Fig. 2, which is defined by the following relations:

$$Z_1 = \frac{h_{01}}{d^\alpha} X + h_{21} X_2 + \mathcal{N}_1,$$
$$Z_2 = h_{02} X + h_{12} X_1 + \mathcal{N}_2,$$
$$Y_1 = h_{03} X + h_{13} X_1 + h_{23} X_2 + \mathcal{N}_3.$$

Define $\mathcal{N}_i$'s to be additive noises, i.i.d. circularly symmetric complex Gaussian RVs with zero-mean; let h_{ij}'s be independent zero-mean circularly symmetric complex Gaussian RVs. d is the random path-loss. The average power of the source and relay inputs X, X_1 and X_2 must not exceed powers P, P_1 and P_2, respectively. Compression is obtained by adding additive noises $\hat{Z}_1 = Z_1 + \hat{\mathcal{N}}_1$, $\hat{Z}_2 = Z_2 + \hat{\mathcal{N}}_2$. It is assumed that the source is not aware of the fading coefficients, the relays know all fading coefficients except h_{i3}'s and the destination is fully aware of everything. The possibilities to choose the proper cooperative strategy are as follows: (i) both relays use DF scheme to transmit the information (full DF case), (ii) both relays use CF scheme to transmit the information (full CF case), where the destination can consider one or both relays as noise to prevent the performance degradation, and (iii) one relay uses DF scheme and the other uses CF scheme (Mixed Coding case). Finally, the relays can select their coding strategy based on the channel parameters (SCS case).

Fig. 3 presents numerical analysis of these strategies. We assume all fading coefficients are of unit variance and so are the noises. d is chosen with uniform distribution between 0 and 0.1, which means the first relay is always around the source. Given this assumption, we suppose that the first relay uses DF in case of mixed coding while the other uses CF scheme. The
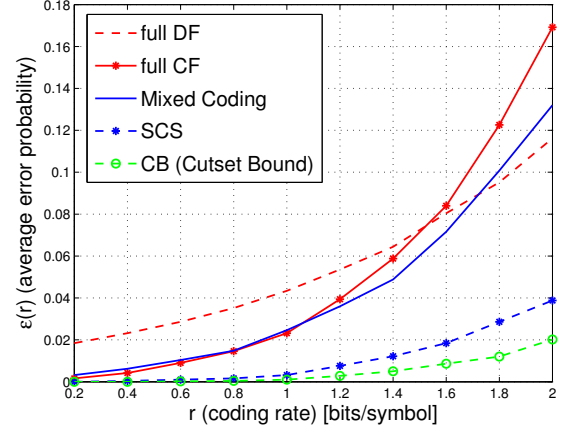

Fig. 3. Asymptotic error probability $\bar{\epsilon}(r)$ vs. coding rate r.

source and relay powers are respectively 1 and 10. It can be seen that none of the non-selective strategies like full DF, full CF and Mixed Coding is not in general the best regardless of fading coefficients. However, if one lets the relay select their strategy given its channel measurement, this SCS will lead to significant improvement compared to the other strategies and becomes close to the cutset bound.

REFERENCES

[1] P. Elias, A. Feinstein, and C. Shannon, "A note on the maximum flow through a network," *Information Theory, IRE Transactions on*, vol. 2, no. 4, pp. 117 –119, 1956.

[2] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *Information Theory, IEEE Transactions on*, vol. 46, no. 4, pp. 1204 –1216, jul 2000.

[3] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *Information Theory, IEEE Transactions on*, vol. 57, no. 5, pp. 3132 –3152, may 2011.

[4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *Information Theory, IEEE Transactions on*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.

[5] J. Wolfowitz, "Simultaneous channels," *Arch. Rat. Mech. Anal.*, vol. 4, pp. 371–386, 1960.

[6] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Information Theory*, vol. 44, pp. 2148–2177, October 1998.

[7] M. Effros, A. Goldsmith, and Y. Liang, "Generalizing capacity: New definitions and capacity theorems for composite channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 7, pp. 3069 –3087, 2010.

[8] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2619 –2692, oct 1998.

[9] M. Katz and S. Shamai, "Cooperative schemes for a source and an occasional nearby relay in wireless networks," *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 5138 –5160, nov. 2009.

[10] A. Behboodi and P. Piantanida, "Broadcasting over the relay channel with oblivious cooperative strategy," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, 29 2010-oct. 1, pp. 1098 –1103.

[11] ——, "On the asymptotic error probability of composite relay channels," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, 2011.

[12] G. Kramer and J. Hou, "On message lengths for noisy network coding," in *IEEE Information Theory Workshop*, October 2011, pp. 430 –431.

[13] X. Wu and L.-L. Xie, "On the optimal compressions in the compress-and-forward relay schemes," Feb. 2011. [Online]. Available: http://arxiv.org/abs/1009.5959